

SAC 12 10

What Is Claimed Is:

1. A method of scheduling data for transmission over a communication link based on priorities assigned to the data, comprising:

5 receiving multiple descriptors at a communication interface device, each of said descriptors describing a data portion having an associated priority;

storing said descriptors in a plurality of memories on said communication interface device, wherein each of said memories is configured to store one or more of said descriptors describing data associated with a predetermined priority;

15 maintaining a dynamic weight for each of said plurality of memories, wherein each said dynamic weight corresponds to a threshold amount of data associated with said predetermined priority; and

servicing said plurality of memories, wherein each said servicing of one of said plurality of memories comprises:

(a) receiving a descriptor from said serviced memory;

(b) retrieving data described by said received descriptor;

(c) scheduling said data for transmission via the communication link;

(d) determining whether an amount of data scheduled during said servicing for transmission via said communication link exceeds said threshold amount of data corresponding to said dynamic weight for said serviced memory; and

20 (e) repeating states (a) through (d) for a next descriptor in said serviced memory if said amount of data scheduled for transmission during said servicing is less than said threshold amount of data.

25 2. The method of claim 1, wherein said servicing further comprises:

(f) determining if said serviced memory contains a descriptor.

3. The method of claim 1, wherein said servicing further comprises:
(f) determining whether a dynamic weight for one of said plurality of memories has changed.

5

4. The method of claim 1, wherein said servicing further comprises:
(f) if said amount of data scheduled for transmission exceeds said threshold amount of data, decreasing said threshold for a next servicing of said serviced memory.

10

5. The method of claim 4, further comprising:
if any of said dynamic weights changes prior to said next servicing, reinstating the pre-decreased threshold for said next servicing.

15

6. The method of claim 1, wherein said receiving multiple descriptors comprises:
determining if a first memory of said plurality of memories contains less than a predetermined number of descriptors, wherein said first memory is configured to store one or more descriptors describing data associated with a first priority;

issuing a request to a host computer, said request identifying said first memory;

receiving a first descriptor describing a first set of data having said first priority.

25

7. The method of claim 6, wherein said first descriptor comprises one or more of:

an identifier of a storage area on said host computer containing said first set of data;

an indicator configured to indicate whether said first set of data is a starting portion of data for a packet; and

5 an indicator configured to indicate whether said first set of data is an ending portion of data for a packet.

8. The method of claim 1, further comprising transmitting said data scheduled for transmission via said communication link before the entire contents of a packet comprising said scheduled data are scheduled for transmission.

9. The method of claim 1, wherein each of said dynamic weights is dynamically modifiable to adjust said threshold amounts of data.

15 10. The method of claim 1, wherein the communication interface device is a network interface circuit and the communication link is a network.

11. A method of scheduling data for transmission over a communication link by servicing, in turn, multiple memories associated with data 20 having different priorities, comprising:

storing in a first memory a first set of descriptors associated with data having a first priority, wherein said first memory has a first dynamic weight corresponding to a first threshold amount of data;

25 storing in a second memory a second set of descriptors associated with data having a second priority, wherein said second memory has a second dynamic weight corresponding to a second threshold amount of data; and

in a first servicing turn of said first memory:

receiving a first descriptor from said first memory;
parsing said first descriptor to identify a first data portion having
said first priority;
5 retrieving said first data portion from a host computer memory;
scheduling said first data portion for transmission onto the
communication link; and
determining whether an amount of first priority data exceeding said
first threshold has, during said first servicing turn, been scheduled for
transmission; and
10 if said first threshold has been exceeded, maintaining a first deficit to
determine how much less than said first threshold of data may be scheduled
during a subsequent servicing turn of said first memory, wherein said first deficit
is initially proportional to said excess.

15 12. The method of claim 11, wherein said first servicing turn further
comprises: determining whether one of said first weight and said second weight
has changed.

13. The method of claim 12, wherein said first deficit is set to zero if
20 one of said first weight and said second weight has changed.

14. The method of claim 11, wherein said first servicing turn further
comprises: determining whether said first memory is empty.

25 15. The method of claim 14, wherein said first servicing turn is
terminated if, during said first servicing turn, either said first memory is
determined to be empty or said amount of first priority data scheduled for

transmission exceeds said first threshold.

16. The method of claim 11, wherein said determining comprises:
incrementing a data counter for each unit of first priority data scheduled
5 during said first servicing turn; and
comparing said data counter to said first threshold.

17. The method of claim 16, wherein said data unit is a byte.

10 18. The method of claim 11, further comprising servicing said second
memory in a second turn, wherein said servicing said second memory comprises:
until at least one of:
said second memory is determined to be empty;
one of said first weight and said second weight change; and
15 an amount of data scheduled during said second turn for
transmission over the communication link exceeds the lesser of said
second threshold and said second threshold minus a second deficit,
wherein said second deficit corresponds to an amount of data by which
said second threshold was exceeded in one or more earlier servicing turns
20 of said second memory;
repeatedly:
receiving from said second memory a second descriptor describing
a second set of data having said second priority;
retrieving said second set of data;
25 scheduling said second set of data for transmission via the
communication link; and
tracking an amount of data scheduled during said second turn by

adding the size of said second set of data to a measure of data previously scheduled during said second turn.

19. The method of claim 18, wherein:

5 said first memory corresponds to data having a highest priority; and if one of said first servicing turn and said second servicing turn terminates because one of said first dynamic weight and said second dynamic weight change, said first memory is the next memory serviced.

10 20. The method of claim 11, wherein the method is performed in a network interface circuit and the communication link is a network.

15 21. The method of claim 11, wherein said first dynamic weight is approximately equal to a maximum packet size of the communication link.

22. The method of claim 11, wherein said second dynamic weight is approximately equal to one.

23. A computer-implemented method of transmitting packets onto a 20 network on the basis of priorities assigned to the packets, comprising:

receiving a first descriptor associated with a first packet having a first priority;

receiving a second descriptor associated with a second packet having a second priority different than said first priority;

25 placing said first descriptor in a first memory on a network interface device, said first memory having a first weight corresponding to said first priority, wherein said first weight corresponds to a first amount of data;

placing said second descriptor in a second memory on said network interface device, said second memory having a second weight corresponding to said second priority, wherein said second weight corresponds to a second amount of data;

5 retrieving from said first memory a first set of descriptors including said first descriptor, said first set of descriptors determined on the basis of said first weight;

Sub A2 10 scheduling said first packet for transmission onto a network connection;
determining whether a size of said first packet exceeds said first amount of data;

retrieving from said second memory a second set of descriptors including said second descriptor, said second set of descriptors determined on the basis of said second weight;

15 scheduling said second packet for transmission onto a network connection;
and
determining whether a size of said second packet exceeds said second amount of data.

24. A computer readable storage medium storing instructions that,
20 when executed by a computer, cause the computer to perform a method of scheduling data for transmission over a communication link by servicing, in turn, multiple memories associated with data having different priorities, the method comprising:

25 storing in a first memory a first set of descriptors associated with data having a first priority, wherein said first memory has a first dynamic weight corresponding to a first threshold amount of data;

storing in a second memory a second set of descriptors associated with

data having a second priority, wherein said second memory has a second dynamic weight corresponding to a second threshold amount of data; and

in a first servicing turn of said first memory:

receiving a first descriptor from said first memory;

5 parsing said first descriptor to identify a first data portion having said first priority;

retrieving said first data portion from a host computer memory;

scheduling said first data portion for transmission onto the communication link; and

determining whether an amount of first priority data exceeding said first threshold has, during said first servicing turn, been scheduled for transmission; and

if said first threshold has been exceeded, maintaining a first deficit to determine how much less than said first threshold of data may be scheduled 15 during a subsequent servicing turn of said first memory, wherein said first deficit is initially proportional to said excess.

25. A communication interface device for transmitting prioritized data over a communication link, comprising:

20 a first memory configured to store a descriptor corresponding to a first packet having a first priority, said first memory being associated with a first weight, wherein said first weight corresponds to a first preferred amount of data to be scheduled, during a first servicing turn of said first memory, for transmission over a communication link;

25 a second memory configured to store a descriptor corresponding to a second packet having a second priority, said second memory being associated with a second weight, wherein said second weight corresponds to a second

preferred amount of data to be scheduled, during a first servicing turn of said second memory, for transmission over said communication link;

a transmission queue into which one of said first packet and said second packet is placed for transmission over a communication link; and

5 an arbiter configured to monitor an amount of data retrieved during said servicing turn in which one of said first packet and said second packet is placed in said transmission queue;

wherein said first weight and said second weight are dynamically adjustable.

26. The communication interface device of claim 25, further comprising a loader configured to retrieve said first packet for placing in said transmission queue during said servicing turn of said first memory.

15 27. The communication interface device of claim 26, wherein said loader is further configured to load a next descriptor for storage in one of said first memory and said second memory.

20 28. The communication interface device of claim 25, wherein said arbiter is further configured to determine whether an amount of data placed in said transmission queue during said first servicing turn of said first memory exceeds said first preferred amount of data to be placed in said transmission queue during said first servicing turn of said first memory.

25 29. The communication interface device of claim 28, wherein said first preferred amount of data is reduced by a deficit for a second servicing round of said first memory.

30. The communication interface device of claim 28, wherein said deficit corresponds to an amount of data, beyond said first preferred amount of data, that is placed in said transmission queue during said first servicing turn.

5

31. The communication interface device of claim 26, further comprising a multiplexer configured to pass said descriptor corresponding to said first packet to said arbiter and said loader during said first servicing turn of one of said first memory and said second memory.

10